How I Think

Benjamin S. Skrainka

University College London
Centre for Microdata Methods and Practice

August 10, 2011

Overview

Discuss thinking as Economist vs. Physicist

- Data
- Model
- Verification & Validation
- Theory vs. Experiment/Applied

Case Study: Finite sample bias in the BLP model

Theory vs. Experiment

Theory and experiment have a different relationship in Physics and Economics:

- In Physics:
  - Tight coupling promotes progress in both
  - Frequent falsification of models via experiment
  - Verification of theoretical predictions
- In Economics:
  - Theorists often work with toy models which are untestable
  - Theory often provides clever mathematical framework to justify our ‘intuition’ about real world behavior
  - Empiricists develop own statistical models

Data : Natural Sciences

In Natural Sciences research often proceeds as follows:

- Write grant
- Build experiment
- Collect data from a controlled experiment (unless you are an Astronomer. . .)
- Write paper
- Repeat

Data : Social Sciences (1)

In Social Sciences data is usually of lower quality and designed for a different purpose:

- Survey data
- Purchase marketing data (very expensive)
- Census data
- Beg or cajole data from some personal connection
- But, many confounding factors, most of which are unknown and unobserved

Economists are similar to Astronomers: you can only observe. . .

Data : Social Sciences (2)

- Natural experiment: Is it truly randomized?
- Sometimes run a controlled experiment:
  - Field experiment
  - In the laboratory
  - Undergraduates most common test subjects
  - Dependent on Human Subjects Committee
  - Does it have external validity?

Poor funding to collect good data to answer important social questions even when real money is at stake!

Data : Social Sciences (3)

Endogeneity – unobserved shocks can affect both LHS and RHS:

1. Measurement error
2. Omitted Variable Bias
3. Simultaneity : an unobserved shock which affects RHS and LHS variables.

- Physics concerned with minimizing 1.
- But, in Economics 2. and 3. often dominate

Example

What if ϵ_{it} contains unobserved Ability?

$$Wage_{it} = \alpha\, School_{it} + \beta\, Class_i + \gamma\, FamilyWealth_{it} + \epsilon_{it}$$

Model Building : Physics

Models are usually governed by tight feedback loop with the real world:[1]

- Iterative refinement to understand one set of laws
- Some ever improving approximation to the truth is knowable
- Driven by explaining observations from natural world
- Everyone is basically refining the same model
- General principles/laws apply to all disciplines, e.g. conservation of energy
- Knowledge is deep like a pine tree

[1] Let's not talk about string theory...

Model Building : Economics (1)

In Economics, model building is part art and part science:

- The world is complex
- We are ignorant
- Many (important) factors are unknowable

Good Economics handles this by

- Choosing a sensible functional form to capture relationships of interest
- Using an error structure which soaks up confounding factors
- Identifing important factors while minimizing potential sources of bias

Model Building : Economics (2)

There are economic 'laws' but many depend on the specific problem and application:

- A good economist is conversant in a much larger number of models
- Models often have narrow applicability
- Some cannot be tested
- 'I did not write this model to be tested. I wrote it to tell a story.'
- Knowledge is more like a dense hedge

Verification

Economics lacks a culture of Verification:

- No credit for replication studies
- Consequently, results are almost never verified
- Huge push to publish single author papers to 'signal' your ability, especially early in career
- Sloppy numerical work in most cases:
  - Code and data are unobtainable
  - No unit tests
  - Few Monte Carlo studies to verify
  - Backlash for discussing or asking about numerical details
- When we mentioned we were going to test our estimation code on Monte Carlo data, a colleague said, 'Oh, you are serious...'

Validation

Some concern about validation:

- Testing for model misspecification
- Usually consists of nesting model in a larger statistical model
- No scenario analysis
- No attempt to explore unknown unknowns
  - Focus is on factors causing unobservable shocks
  - Little attempt to determine magnitudes of different unobservables
- Avoidance of anything outside standard toolbox: e.g. preference for pMC over SGI quadrature rules

Case Study: BLP

The Berry, Levinsohn, and Pakes (1995) model of differentiated products dominates the field of Industrial Organization:

- Justification of model rests on one paper
  - Only considers a special case
  - Only uses tools in mainstream Economics toolbox
- Drives $\sim 70\%$ of papers in applied Industrial Organization
- Will soon be used for anti-trust inquiries...

‘The Death Star’

Skrainka (2011) develops an infrastructure to test BLP and other econometric models:

- Prior simulation experiments only look at data with one market
- First simulation study of bias for range of (more) realistic datasets
- Rigorous generation of synthetic data as well as estimation: most econometric theorists use antediluvian numerical methods
- Automated everything possible to avoid stupid errors
- Developed scripts (bash, Python, R) to verify each run
- One of the largest simulations in Economics – but only about $90,000$ CPU-hours

Data Generation

Generating realistic data is harder than estimating the model:

- Used structural model so data has correct statistical properties as number of markets, T, and products, J, $\to \infty$
- Smaller datasets are subsets of larger datasets
- Generated prices from Bertrand-Nash equilibrium:
  - Must find root of nonlinear set of equations which often mislead a solver
  - Has T dense $J \times J$ blocks
  - Used Path 5.0.00
- Similar technology to estimation (Eigen, C++, etc.)
- Generated up to 50 markets and 100 products

Data Estimation

Used current best methods for estimation

- Su & Judd (2011): Bi-level optimization + SNOPT solver
- Skrainka & Judd (2011): Sparse grids quadrature rules
- Eigen, C++, higher precision (long double)
- 50 starts to hopefully find global min
- Restarts to ensure found global min
- Used two different instrumentation strategies

Parallelization

Very easy to parallelize:

- Each replication and start is independent
- Use Parameter Sweep
- Very easy with PBS
- Compute results based on best start

Bias: Point Estimates

T	J	Bias	Mean Abs Dev	Med Abs Dev	RMSE
1	12	−2	3	1.3	5.7
1	24	−0.72	1.9	1.2	3.2
1	48	−0.52	1.9	1.2	3
1	100	−0.57	1.7	1.3	2.3
10	12	−1.7	2.6	1.1	6
10	24	−0.65	2	1.3	3.6
10	48	−0.64	1.9	1.3	3.2
10	100	−0.83	2	1	3.9
25	12	−0.62	1.9	1.2	3.1
25	24	−0.96	2.3	1.4	3.7
25	48	−1.3	2.8	1.2	7.6
25	100	−0.95	2.1	1.1	3.7
50	12	−0.39	1.6	1.1	2.7
50	24	−1.2	2.5	1.1	5.4
50	48	−1.2	2.2	1.3	5.2
50	100	−0.63	1.9	1.3	3

Bias: Own-Price Elasticities

T	J	Bias	Mean Abs Dev	Med Abs Dev	RMSE
1	12	−0.77	2.2	0.94	4.9
1	24	−0.095	1.5	0.77	3.3
1	48	−0.082	1.6	0.91	2.7
1	100	−0.39	1.5	0.98	2.5
10	12	−0.5	1.7	0.81	3.3
10	24	−0.57	1.7	0.83	3.3
10	48	−0.16	1.5	0.97	2.2
10	100	−0.53	1.7	0.93	3.3
25	12	−0.3	1.4	0.94	2.7
25	24	−0.72	1.8	1.1	3
25	48	−0.87	2.2	1.1	4.9
25	100	−0.61	1.7	0.97	2.7
50	12	−0.43	1.5	0.94	2.6
50	24	−0.77	1.9	0.91	3.8
50	48	−0.97	1.9	1.1	4
50	100	−0.56	1.8	1.1	2.9

Bias: Quadrature

	Bias		Mean Abs Dev		Med Abs Dev		RMSE	
	SGI	pMC	SGI	pMC	SGI	pMC	SGI	pMC
θ_{11}	0.96	12.34	2.29	13.25	1.20	3.64	4.00	28.92
θ_{12}	0.02	−0.13	0.52	0.38	0.22	0.33	0.94	0.48
θ_{13}	−0.28	−0.38	1.47	1.21	0.62	0.99	3.01	1.51
θ_{21}	22.57	128.22	23.01	128.24	2.62	34.06	81.76	253.87
θ_{22}	0.02	−0.04	0.12	0.16	0.07	0.13	0.19	0.20
θ_{23}	0.08	0.64	0.36	0.75	0.16	0.79	0.75	0.90

Table: Comparison of bias in point estimates : SGI vs. pMC for T=2 markets and J=24 products with 165 nodes.

Impact on Policy

This study will hopefully impact Econometrics and applied work driving policy:

- Asymptotic inference is not valid with this model
- Possible to test Econometric tools/estimators more rigorously than previously attempted
- BLP is poised to become a key model for anti-trust investigation but estimation strategy is a failure:
  - Traditional strategy is extremely biased
  - Best strategy is also biased!
  - Lack of control for endogeneity for typical datasets
- Conventional wisdom that increasing the number of markets will improve results is wrong because most variation is with-in market!

Conclusion

Economists can do much more verification and validation:

- Modern resources make it easy to run large scale Monte Carlo experiments:
  - Test estimators
  - Investigate more scenarios and counter-factual policy experiments
  - Richer policy analysis
- BLP model is seriously biased:
  - Need a new model
  - Need better identification strategy
- Other models probably have similar problems, but no one has attempted to validate them
- Better data and methods needed to provide better answers to important questions